Media Forensics (MediFor)

Ensuring "seeing is believing"

Dr. David Doermann

I2O

October 2, 2015

MediFor Proposers Day Agenda

Start	End	Session
09:00	10:00	Registration
10:00	10:10	Security Briefing
10:10	10:15	Welcoming Comments
10:15	10:30	CMO Briefing (Attendees submit questions)
10:30	11:30	Program Discussion (Attendees submit questions)
11:30	12:30	Attendees may speak for 2 minutes for teaming purposes. No slides or handouts. DARPA representatives will not be present.
12:30	01:30	Lunch – on your own
01:30	02:30	Question & Answer Session

Media Forensics (MediFor) Program
DARPA-BAA-15-58

Mark Jones

DARPA Contracts Management Office

Proposers Day
Arlington, VA
October 2, 2015

DISCLAIMER

If the BAA contradicts any information in these slides,

the <u>BAA</u> takes precedence.

BAA OVERVIEW

BAA follows procedures in accordance with FAR 35.016.

BAA is posted on FEDBIZOPPS at www.fbo.gov and Grants.gov at http://www.grants.gov/ (as well as any future amendments).

Proposals due by 12:00 noon ET on November 24, 2015

BAA covers all info needed to submit proposals. Follow instructions for proposal preparation and submittal.

The BAA FAQ and slides from this meeting will be posted to http://www.darpa.mil/work-with-us/opportunities/darpa-baa-15-77.

BAA ELIGIBILITY

All interested/qualified sources may respond subject to the parameters outlined in the BAA.

Foreign organization/individuals – check all applicable Security Regulations, Export Control Laws, Non-Disclosure Agreements, and any applicable governing statutes.

FFRDCs and Government entities
- Subject to applicable direct competition limitations
- Must clearly demonstrate eligibility per BAA

Real and/or Perceived Conflicts of Interest
- Identify any conflict
- Include mitigation plan

TA2 Only - personnel cleared at a minimum to TS who are eligible for SCI

POTENTIAL AWARD INFORMATION

Three Technical Areas (TAs) – TA 1, 2 and 3

TA1 Subareas: TA1.1, TA1.2 and TA1.3

Anticipate multiple awards – exact award numbers or amounts have not been predetermined

TA2 workload anticipated less in Phase 1, but increase as program progresses

Program structured in 3 phases – Base and 2 Options

Abstract Phase

Award may be procurement contract, cooperative agreement or Other Transaction Agreement (OTA).   No grants will be awarded.

ABSTRACTS

Due by 12:00 noon (ET), October 9, 2015

Four page limit – do not include any classified information

Submit abstracts via DARPA's web portal – do NOT submit via Grants.gov

DARPA reply will provide encourage/discourage feedback – discourage will include detailed rationale of decision

Abstract submissions are not a prerequisite for proposal submissions

DARPA will review all properly submitted proposals without regard to feedback

DO NOT wait until the last minute to submit abstracts – the submission deadline as outlined in the BAA will be strictly enforced

PROPOSAL PREPARATION INFORMATION

Proposals consist of two volumes – Technical and Cost.

Volume 1 - Technical and Management
- Volume 1 has maximum page limitations specific to TAs being proposed
- Includes <u>mandatory</u> Appendix A – does not count towards page limit.
- Includes optional Appendix B (TA1 Only) – does not count towards page limit
- Includes optional Appendix C (TA2 Only) – <u>does</u> count towards page limit
- Includes optional Appendix D – does not count towards page limit

Volume 2 – Cost - No page limit.

The BAA will describe the necessary information to address in each volume –
- Make sure to include every section identified.
- If a section does not apply – put "None" (e.g., Animal Use – None, OCI - None)
- Include a <u>working/unprotected</u> spreadsheet as part of your Cost Volume submission.
- Review individual TA descriptions, IP and the deliverables section for submittal information

PROPOSAL PREPARATION INFORMATION

Can submit proposal for any number of TAs

Each proposal must only address <u>one</u> TA

Proposals for TA1 subareas can address any number of subareas under a single proposal

Conflicts of interest between TA1/2 vs TA3 – If single entity submits proposals against multiple TAs, this may create a conflict that would be resolved at the Government's discretion.

Program Stresses Open Exchange of Information – will utilize Associate Contractor Agreement contract clause (or similar condition in non contract awards)

PROPOSAL PREPARATION TIPS

Statement of Work (SOW) – Write a SOW as if it were an attachment to a contract

- o Don't use proposal language (e.g. we propose to do . . .)

- o Break out work between any phases/time periods identified in the BAA

- o Succinctly and clearly define tasks & subtasks

- o Do not include any proprietary or classified information!

- **Risk** – Do not be afraid to address Risk in Technical Volume

  - o Identify risk(s) to show an understanding of technical challenge(s)

  - o Discuss potential mitigation plans / alternative directions

PROPOSAL PREP – INTELLECTUAL PROPERTY RIGHTS

Government desires, at a minimum, **Government Purpose Rights** for any proposed <u>noncommercial</u> software and technical data. (SEE DFARS 227 for Patent, Data, and Copyrights)

Since MediFor will emphasize creating and leveraging open architecture technology, IP rights and software licenses asserted by proposers are strongly encouraged to be aligned with this goal.

Data Rights Assertions – IF asserting **less than Unlimited Rights**:

- Provide and justify basis of assertions

- Explain how the Government will be able to reach its program goals (including transition) within the proprietary model offered; and

- Provide possible nonproprietary alternatives

IF proposed solution utilizes commercial IP – submit copies of license with proposal

ITEMS TO NOTE

Work expected to be fundamental research

Understand and comply with SAM, E-verify, FAPIIS, i-Edison and WAWF. Links are found in the BAA.

For planning purposes - anticipate Program Start Date as May 1, 2016

Subcontracting Issues

- Non-Small Businesses: Subcontracting Plans required for FAR-based <u>contracts</u> expected to exceed the applicable threshold.
- Subcontractor cost - Proposals must include, at a minimum, a non-proprietary, subcontractor proposal for EACH subcontractor.
- If utilizing FFRDC, Government entity, or a foreign-owned firm as a subcontractor, submit their required eligibility information, as applicable.

ITEMS TO NOTE CONTINUED

Proposals must be valid for a minimum of 120 days

If a prospective proposer believes a conflict of interest exists or has a question on what constitutes a conflict - promptly raise the issue with DARPA

Document files must be in .pdf, .odx, .doc, .docx, .xls, and/or .xlsx formats.

Submissions must be written in English.

PROPOSAL SUBMISSION

Submissions will be UNCLASSIFIED with the exception of Appendix C. Appendix C must be received before the BAA proposal submission dead line.

Follow submission procedures outlined in the BAA. DO NOT submit proposals except as outlined in the BAA (e.g., email/fax submissions will NOT be accepted).

Proposals for Cooperative Agreements will utilize the Grants.gov website for uploading proposals.

Proposals for Procurement Contracts/OTAs will utilize DARPA's web-based upload system:

- o If not previously registered – 2 step registration process
- o Submission must be in a single zip file not exceeding 50 MB
- o When submitting – make sure to drop files in correct BAA
- o Must **FINALIZE** submission prior to closing to be considered

DO NOT include any classified information in the unclassified portion of the proposal or it may be deemed non conforming.

DO NOT wait until the last minute to submit proposals – the submission deadlines as outlined in the BAA will be strictly enforced

EVALUATION / AWARD

No common Statement of Work - Proposal evaluated on individual merit and relevance as it relates to the stated research goals/objectives

Evaluation Criteria (listed in descending order of importance) are: (a) Overall Scientific and Technical Merit; (b) Potential Contribution and Relevance to the DARPA Mission; and (c) Cost Realism. For TA2 ONLY, (d) Plans and Capability to Accomplish Technology Transition.

Evaluation done by scientific/technical review process. DARPA SETAs with NDAs may assist in process.

Government reserves the right to select for award all, some, or none of the proposals received, to award portions of a proposal, and to award with or without discussions.

COMMUNICATION

Prior to Receipt of Proposals – No restrictions, however Gov't (PM/PCO) shall not dictate solutions or transfer technology. Unclassified FAQs will be periodically posted to this BAA's DARPA web page.

After Receipt of Proposals – Prior to Selection: Limited to PCO – typical communication to address proposal clarifications.

After Selection/Prior to Award: Communications range from technical clarifications/revisions to formal cost negotiations. May involve technical as well as contracting staff.

Informal feedback for proposals not selected for funding may be provided once the selection(s), if any, are made.

Only a duly authorized Contracting Officer may obligate the Government

TAKE AWAY

Submit abstracts/proposals before the due date/time - Do NOT wait until the last minute to submit.

Read and understand the BAA - Follow the BAA when preparing abstracts and proposals.

Be familiar with Government IP terms from the DFARS Part 227.

Submit <u>working/unprotected</u> spreadsheet(s).

The Contracting Officer is the only Government official authorized to obligate the Government.

Media Forensics
(MediFor)

Ensuring "seeing is believing"

Dr. David Doermann

I2O

October 2, 2015

MediFor Proposers Day Logistics

- BAA Location and Dates
  - Posted on FedBizOpps website (http://www.fedbizopps.gov)
    and Grants.gov website (http://www.grants.gov)
  - Posting Date: September 29, 2015
  - Abstract Due Date: October 9, 2015, 12:00 noon (ET)
  - BAA Closing (Proposal Due Date): November 24, 2015, 12:00 noon (ET)
- Procedure for Questions/Answers Today
  - Questions can be submitted until 11:30am to MediFor@darpa.mil or on 3x5 cards
  - Questions will be answered during Q&A session in the afternoon
  - Waiting until the session is complete is encouraged to avoid repetition
- Websites
  - Proposers Day website
  - MediFor program website
  - The BAA FAQ and slides from this meeting will be posted to
    http://www.darpa.mil/work-with-us/opportunities/darpa-baa-15-77.

MediFor Program Manager Presentation

- Program Overview

  - Problem Statement

  - Program Goals

- Program Structure
  - Technical Area 1:   Integrity Analytics Research and Development
  - Technical Area 2:   Integrity Reasoning Engine and MediFor Console
                         Development
  - Technical Area 3:   Media Resource Development: Corpora Creation,
                         Manipulation, and Annotations
  - Evaluation

- Expected Program Schedule

- Proposal Information

Visual Media Assets

Image Intelligence Applications

Analysis

Investigation

Prosecution

Diverse sources

"Expert" Manipulation

Ad-hoc, manual integrity assessment

Growth in Visual Content in Social Media

Assets per Day (Billions)

2.1
1.8
1.5
1.2
0.9
0.6
0.3
0

2008 2009 2010 2011 2012 2013 2014

Software Imbalance

Software Products

10,000

1,000 — **2174**

100

10 — **9**

1

Manipulation
Detection

2008 2009 2010 2011 2012 2013 2014

Meeker Internet Report 2014
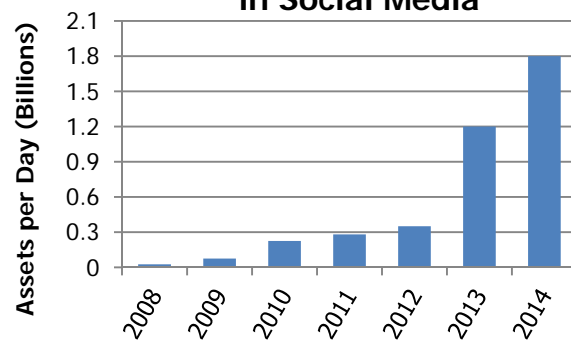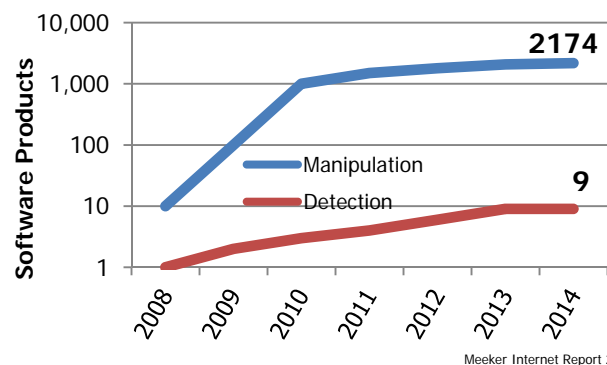
Current Capabilities

- No measure of integrity currently exists.

- Analysts capable of generating few authentication reports/day

22

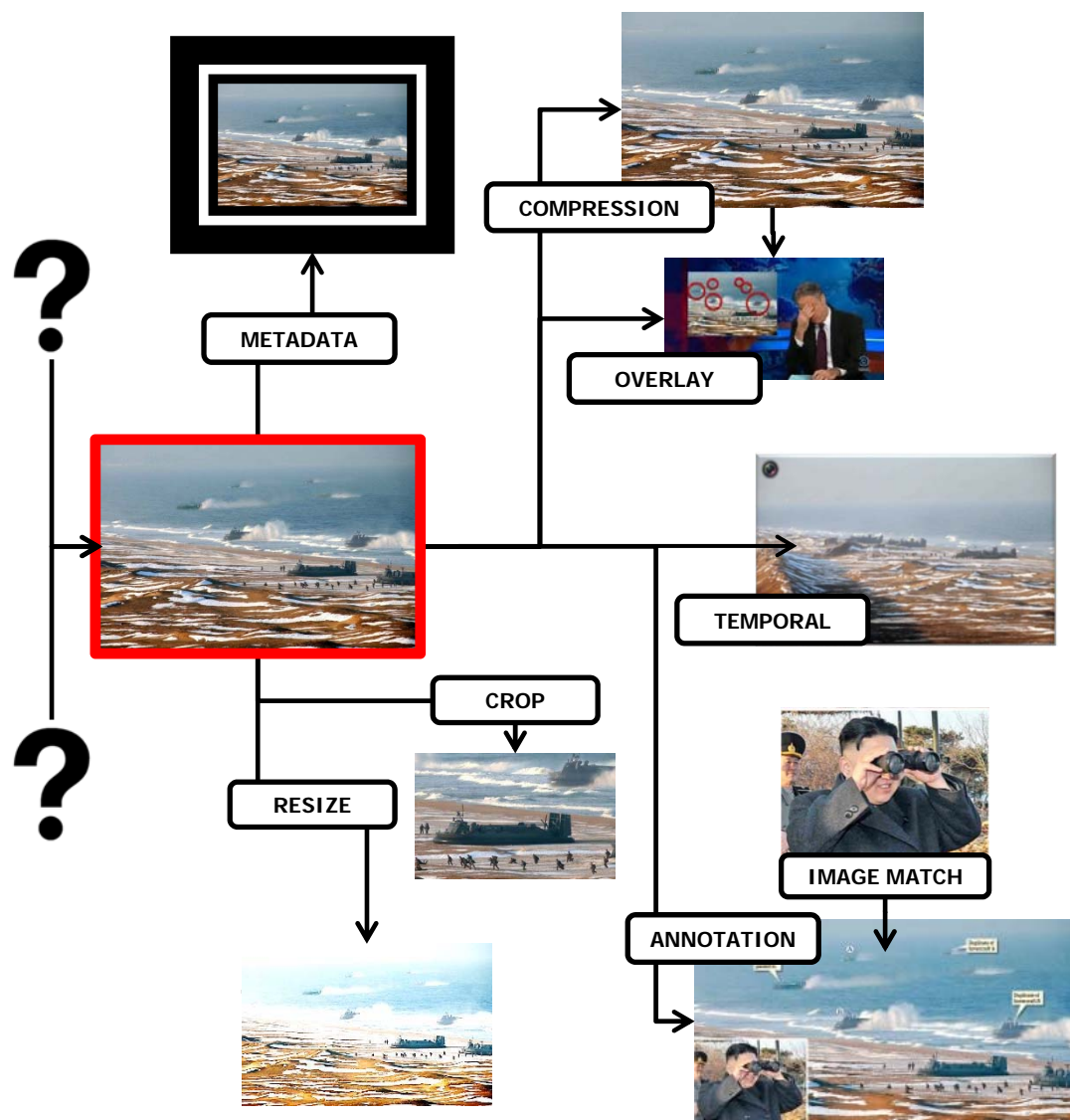# What is the Integrity of This Image?

How Can Assets Be "Related"

Digital Image/Video genealogical history

- Source
- Imaging Device
- Format
- Editing Tools
- Manipulations
- Image Content
- Viewpoint
- Scene/Event
- Temporal Order
- ...

The Essential Elements of Media Integrity

Digital Integrity

Are the pixels/representations inconsistent?

Blurred edges, replicated pixels, mangled compression?

Physical Integrity

Are the laws of physics violated?

Inconsistent shadows and lighting, elongation/compression, multiple vanishing points?

Semantic Integrity

Is a hypothesis about a visual asset disputable?

Contradicting evidence in associated images or resources. Was the image repurposed? *Dates/times* inaccurate?
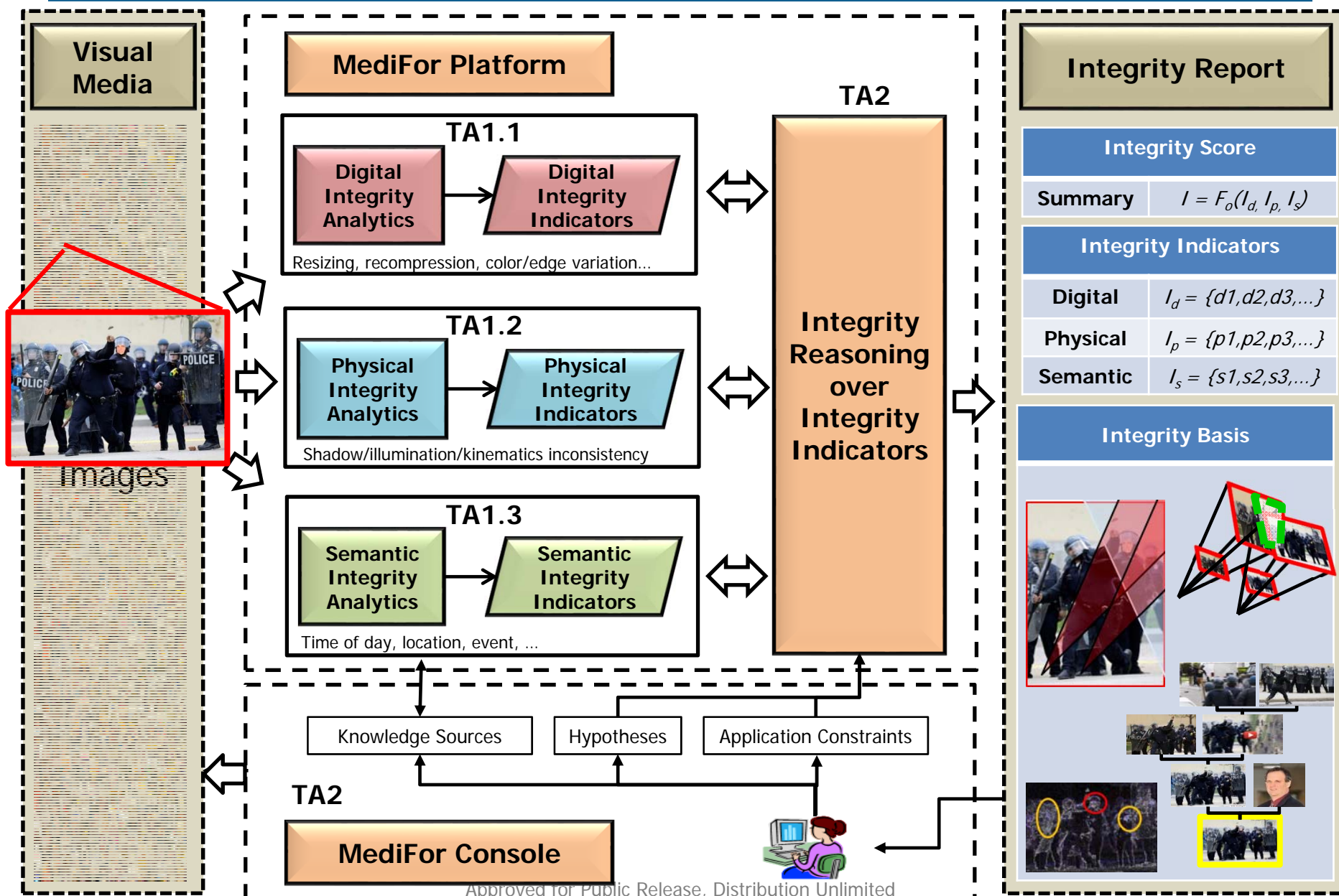
Integrity Reasoning

How can one assimilate digital, physical and semantic integrity indicators into an integrated assessment?

Unifying indicators from integrity analytics, domain knowledge, and application constraints.

Visual Media

Images

MediFor Platform

TA1.1

Digital Integrity Analytics → Digital Integrity Indicators

Resizing, recompression, color/edge variation...

TA1.2

Physical Integrity Analytics → Physical Integrity Indicators

Shadow/illumination/kinematics inconsistency

TA1.3

Semantic Integrity Analytics → Semantic Integrity Indicators

Time of day, location, event, ...

TA2

Integrity Reasoning over Integrity Indicators

Knowledge Sources | Hypotheses | Application Constraints

TA2

MediFor Console

Integrity Report

Integrity Score	
Summary	$I = F_o(I_d, I_p, I_s)$

Integrity Indicators	
Digital	$I_d = \{d1, d2, d3, ...\}$
Physical	$I_p = \{p1, p2, p3, ...\}$
Semantic	$I_s = \{s1, s2, s3, ...\}$

Integrity Basis

Digital Integrity

Physical Integrity
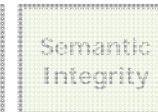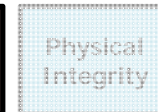
Semantic Integrity

Are the pixels/representations consistent or inconsistent? – Blurred edges, replicated pixels, mangled compression?

State of the Art
- Manipulations are not classified or characterized in detail
- Parameter and process instantiation often done manually
- Little attempt to combine evidence from multiple sources

Primary Challenges
- Classifying as real, computer generated, or a composite
- Gathering evidence and methods of manipulation/alteration
- Determining origin of the material used in editing
- Localization of manipulations
- Automated and efficient analytics that are robust to compression, degradation and size

COLOR SHIFTING

COPY-PASTE

COMPOSITING

Are the laws of physics violated? – Inconsistent shadows and lighting, elongation/compression, multiple vanishing points?

State of the Art
- Automated technology does not exist or is very brittle
- Manual processes are extremely time consuming
- Performance is not robust to simple quality shifts and image transformations

REFLECTION ANALYSIS

Primary Challenges
- Automatic detection of artifacts such as shadows
- Consistency of effects such as shadows and reflections
- Consistency of physical and kinematic properties
- Robustness to simple shifts in degradation, lighting, vantage point

SCENE DYNAMICS
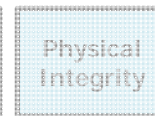

SHADOWS AND HIGHLIGHTS

Are hypotheses about a visual asset supported/contradicted? – Is there contradicting evidence in associated images? Was the media asset repurposed? Are dates/times/locations verifiable?
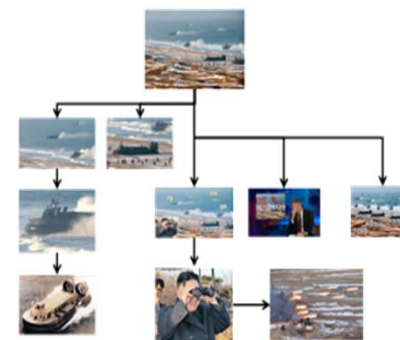
State of the Art
- Association and classification work well when there is dense coverage and unique features (text, unique objects), but fail otherwise
- Techniques have low detection and high false alarm rates
- Verification requires search over large knowledge bases
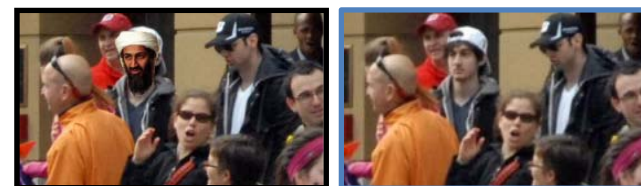
Primary Challenges
- Verifying spatial temporal assumptions
- Uncovering the provenance of an asset or event
  - Search for other assets which are ancestors, decedents, or related
  - Determine which device(s) acquired this asset(s)
- Analyzing and detecting spatio-temporal evolution of events
- Dealing with sparse temporal and spatial coverage at very large scales
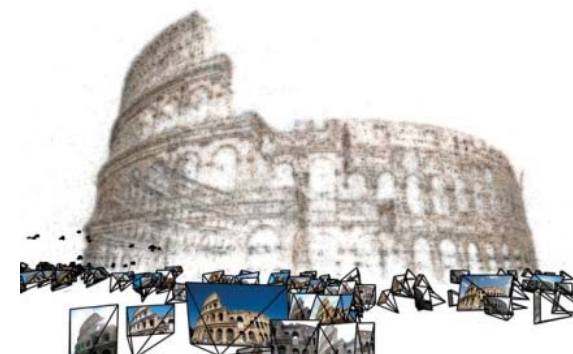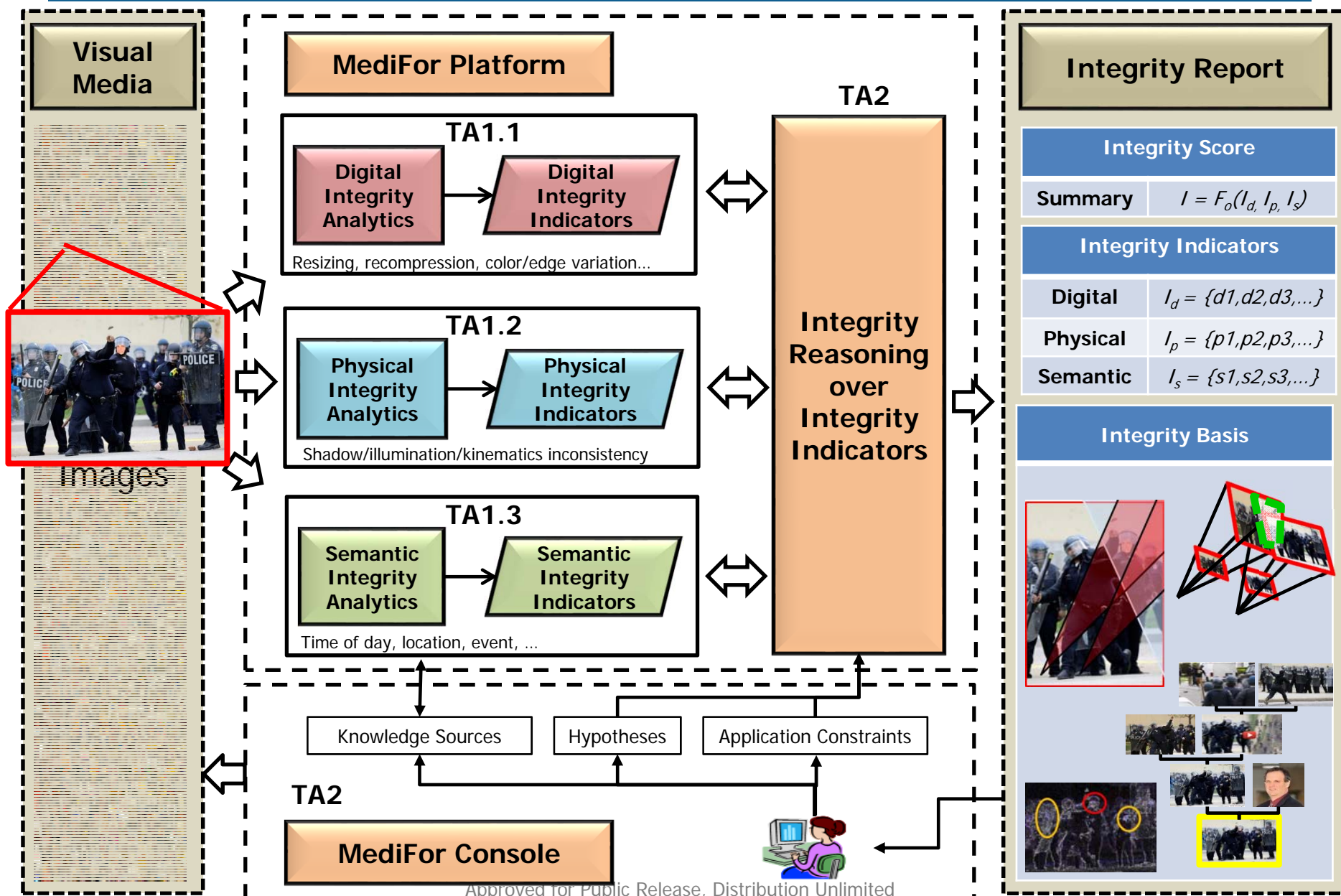
DENSE 3D RECONSTRUCTION

MEDIA PROVENANCE

INFORMATION PROPAGATION

MediFor Platform

TA1.1

Digital Integrity Analytics → Digital Integrity Indicators

Resizing, recompression, color/edge variation...

TA1.2

Physical Integrity Analytics → Physical Integrity Indicators

Shadow/illumination/kinematics inconsistency

TA1.3

Semantic Integrity Analytics → Semantic Integrity Indicators

Time of day, location, event, ...

TA2

Integrity Reasoning over Integrity Indicators

Visual Media

Images

Knowledge Sources | Hypotheses | Application Constraints

TA2

MediFor Console

Integrity Report

Integrity Score	
Summary	$I = F_o(I_d, I_p, I_s)$

Integrity Indicators	
Digital	$I_d = \{d1, d2, d3, ...\}$
Physical	$I_p = \{p1, p2, p3, ...\}$
Semantic	$I_s = \{s1, s2, s3, ...\}$

Integrity Basis

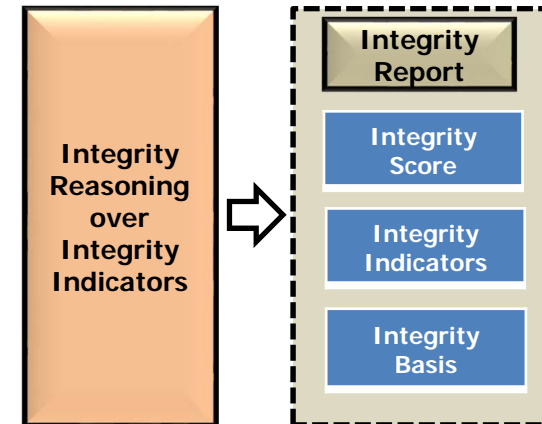# TA2: Integrity Reasoning - Scoring

How can one assimilate digital, physical and semantic integrity indicators into an integrated assessment? – Unifying indicators from integrity analytics, domain knowledge, and application constraints

State of the Art
- Manual processes which are application dependent
- Subjective measures of authenticity
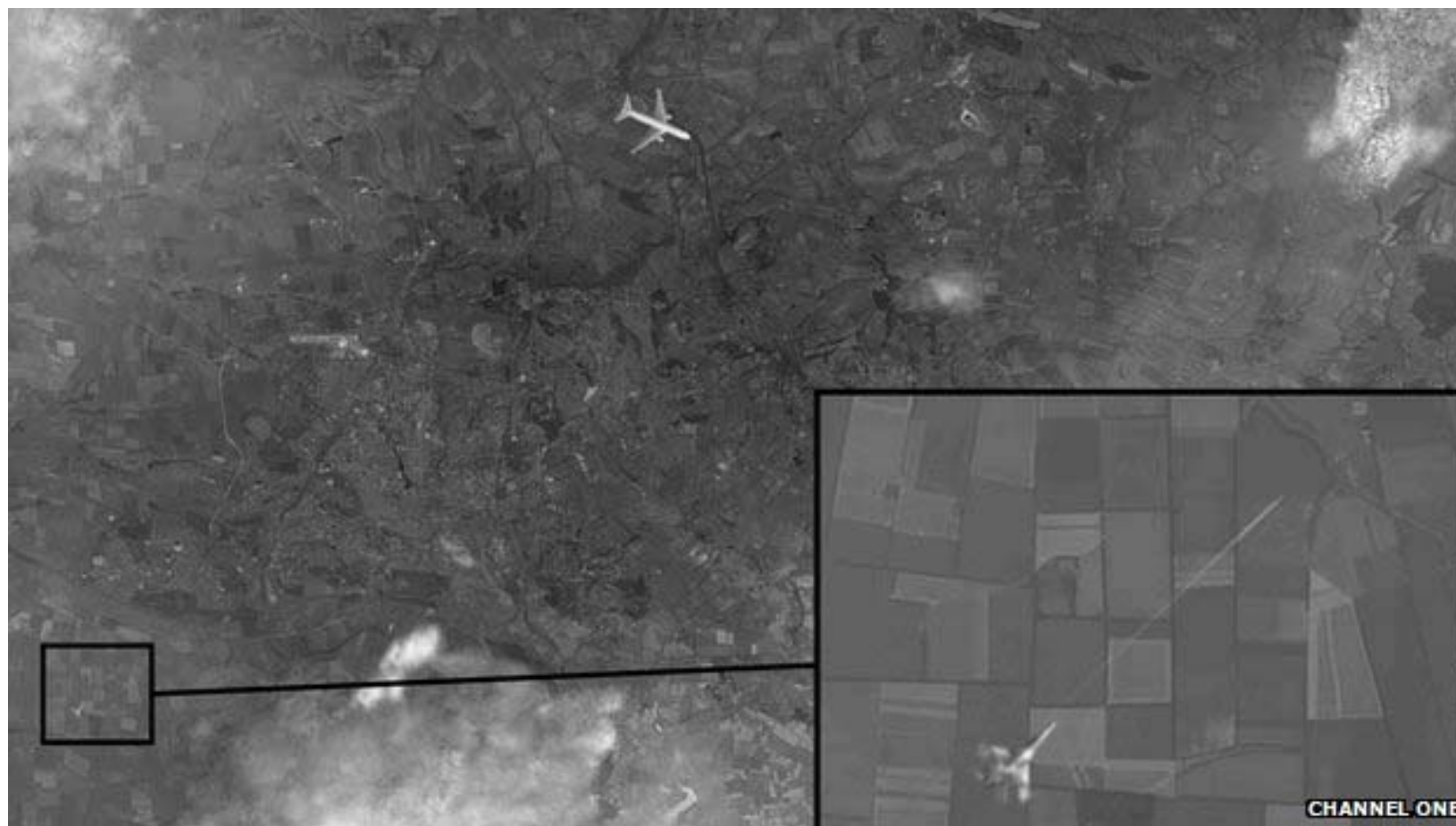- No quantitative standard exists

Primary Challenges
- Establishing a confidence for various analytics
- Combining scores of the different integrity modules
- Developing
  - a score that is generic
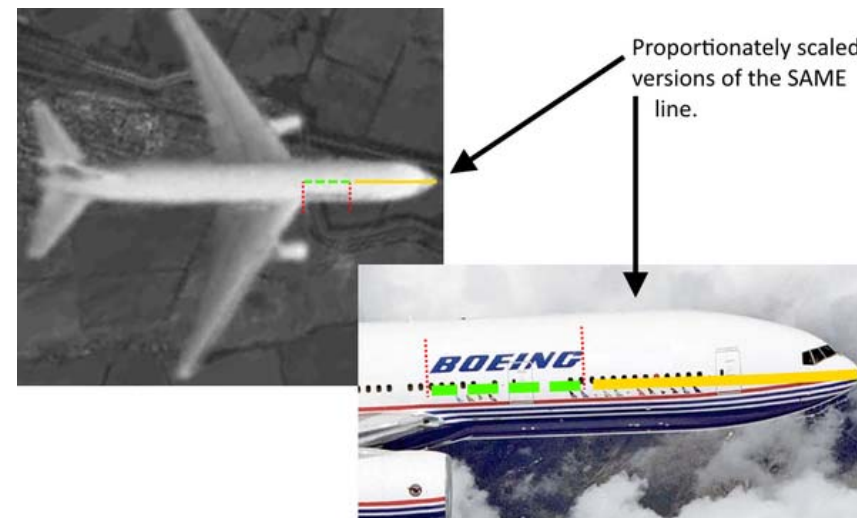 - a representation that can be mapped to a specific application

Integrity Reasoning over Integrity Indicators

Integrity Report

Integrity Score

Integrity Indicators

Integrity Basis
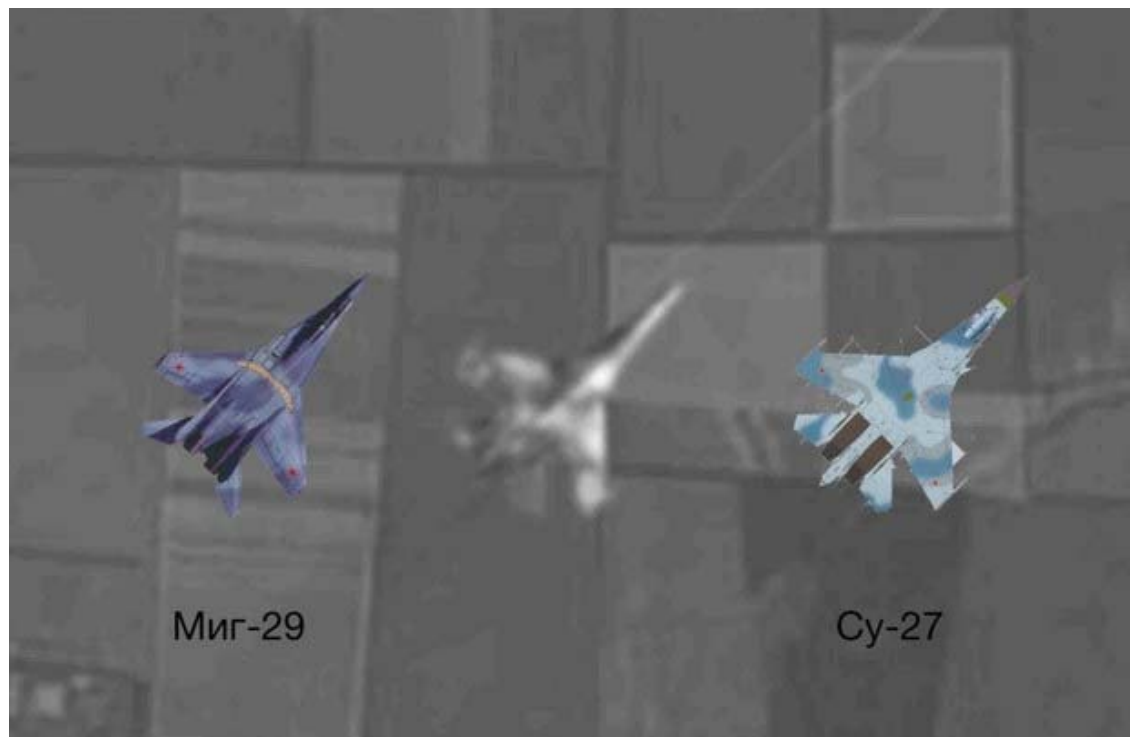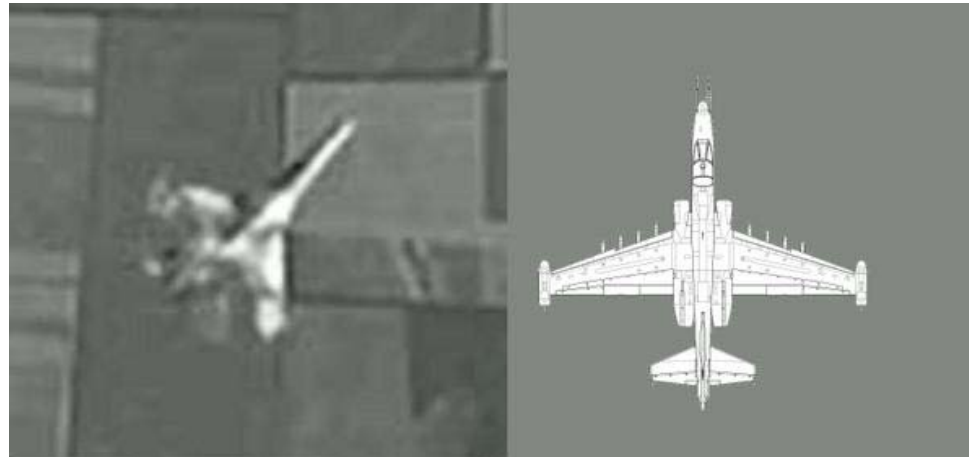
Part of Image Came From Google Cache in 2012

Image compressed & EXIF removed

Cropped

Resized

Copy / Paste

Delete

CGI overlay

EXIF:

File name : img_1771.jpg
File size : 43272641 Bytes
MIME type : image/jpeg
Image size : 3000 x 2500
Camera make : Canon
Camera model : Canon PowerShot S40
Image timestamp : 2003:12:14 12:01:44

EXIF:

Blank

	Original	70 Jpeg Quality	Resized 50%	No Exif	Copy / Paste	CGI Overlay	Cropped	Delete
Average "Joe"	🟢	🟢	🟢	🟢	🟢	🟡	🟡	🔴
Investigator	🟢	🟢	🟡	🟡	🟡	🔴	🔴	🔴
Prosecutor	🟢	🟡	🟡	🔴	🔴	🔴	🔴	🔴

## What is the role of the MediFor Console?

- Allow analysts / operators systems configuration and integrity report access

    - Integrity Report = Integrity Score + Basis

- Provide access to external knowledge sources

    - i.e. Manufacturer color/compression tables, Knowledge of weather/climate

- Communicate analyst provided hypotheses

    - Limited structured / templated formats

    - i.e "Time = 7:00:00 December 21, 2015"

    - **Not** "Is the plane shadow authentic"
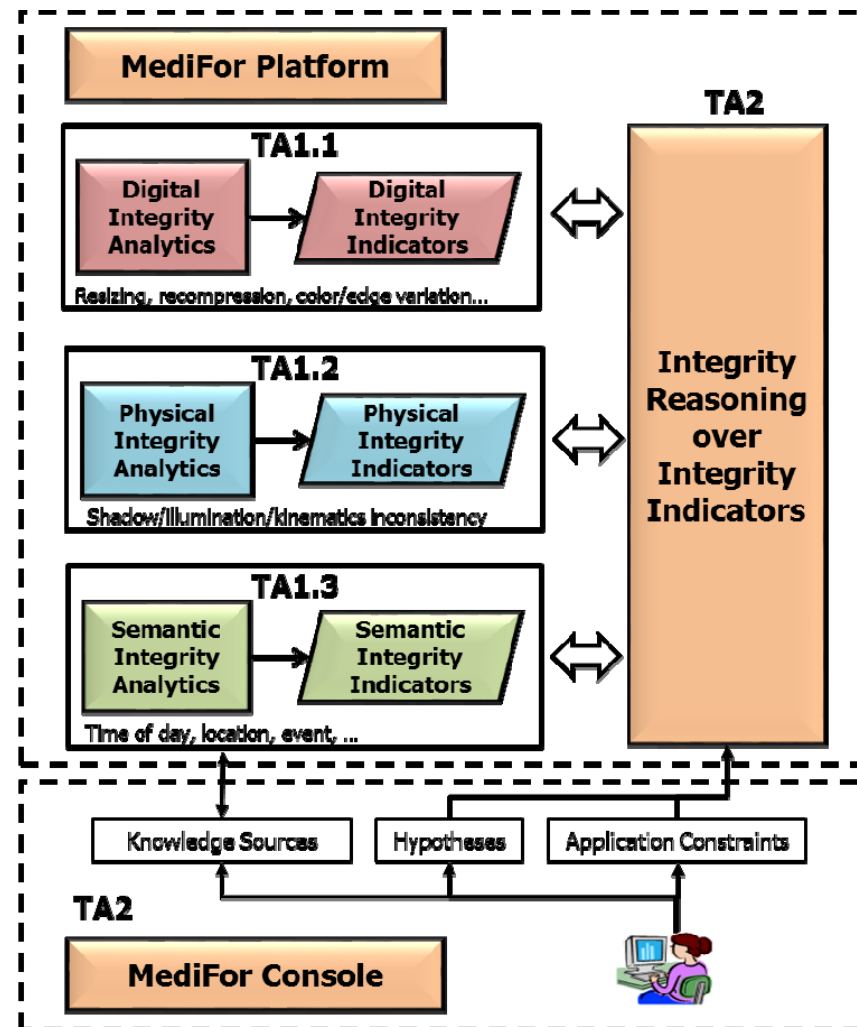
- Accept and share application constraints
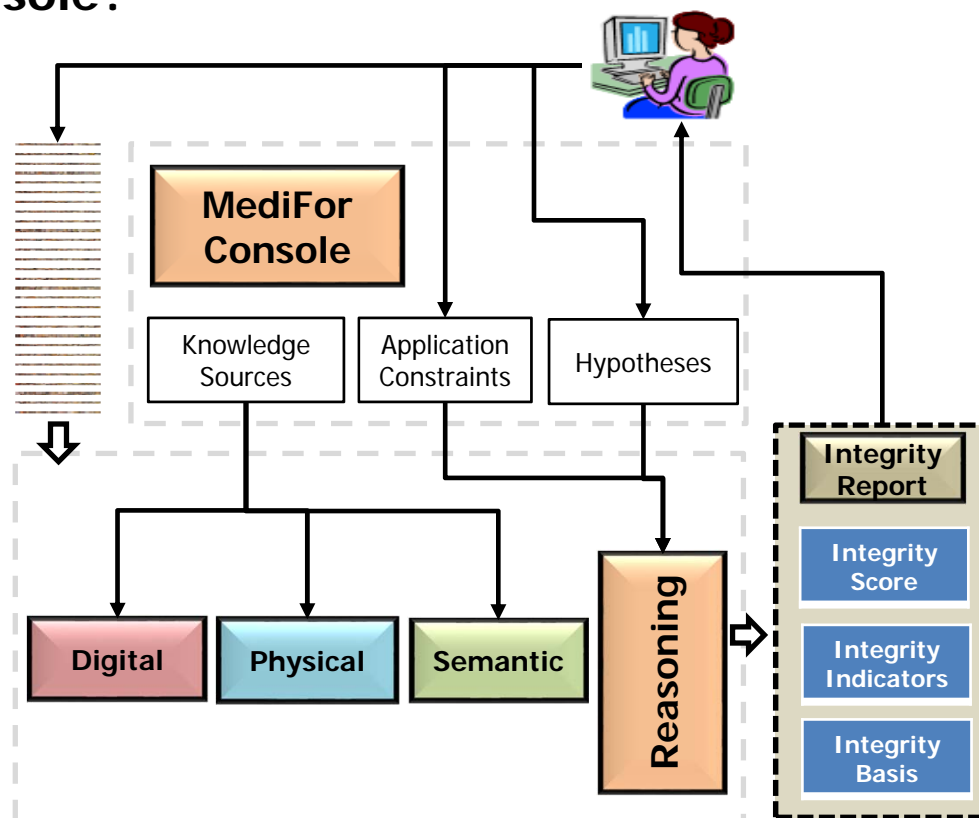
The console will **not**:

- Provide on-demand interaction with individual images

- **Architecture Goals**
  - Open
  - Standards-based
  - Plug-and-play
  - Scalable
  - Ease of interoperability and integration

- **MediFor API**
  - Interface between TA1 analytics and TA2 reasoning
 - Should not constrain TA1 teams

Data & Manipulations

- Data gathered from public sources
- Data divided into *world* and *probe* sets
- Manipulations, trace evidence and event scenarios will be *embedded or contributed*
- Manipulations with varying levels of integrity
- Complexity will escalate across phases.

Asset and Anomaly Volumes

Phase 1   Phase 2   Phase 3

## Example Manipulations and Adversarial Scenarios

- Copy / Paste
- Eraser / Inpainting
- Cropping
- Seam Carving
- Metadata Forgery
- Retouching
- Multiple Compression
- Artificial shadows
- Artificial light sources
- Resizing
- Blur / Sharpening
- Contrast
- Rotation
- Median Smoothing
- Lighten / darken
- Text Overlays
- Warping
- Color Adjustment

- Temporal Reordering
- Repurposing
- Anti-Forensics
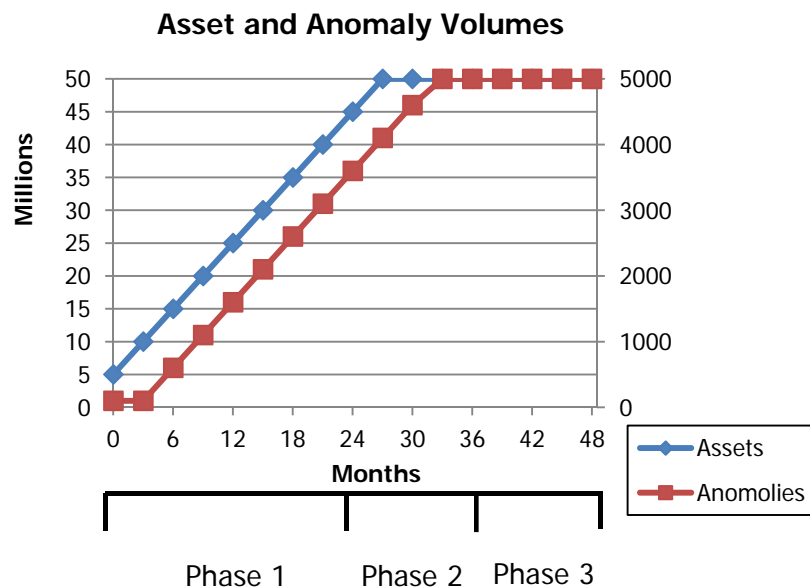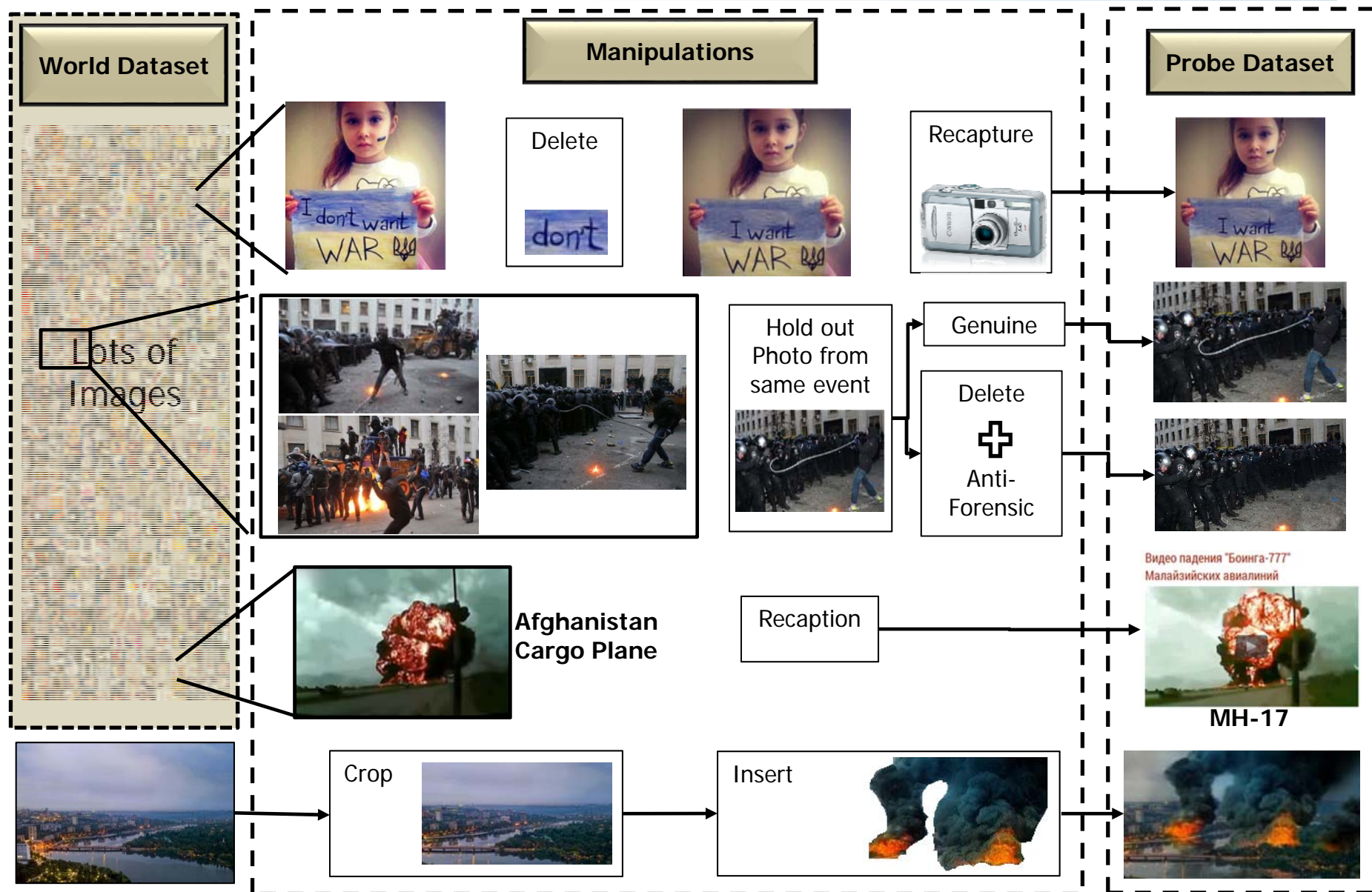
MediFor Evaluation

MediFor evaluations will drive program goals, stimulate community interest, and engage image analysts stakeholders

NIST: Technical Evaluations

NIST: Open Challenges

Integrity Analytics Evaluation (Manipulation)
- Determine if asset genuine/edited/manipulated

Association Evaluation (Provenance)
- Map the who, what, and when (genealogy) of an asset

Verification Evaluation (Hypotheses)
- Given a claim about an asset, prove, disprove, and/or identify inconsistencies based on external resources and document

System Evaluation (Phases 2 & 3)
- Accuracy and completeness of integrity report
- System performance

Field Evaluation (Phases 2 & 3)
- Real world evaluation run with DOD partners

MediFor Portal
- NIST will provide a community-wide portal to support evaluations and challenges
- Challenges will be held annually then remain open with an active leaderboard

FoolMyMediFor
- Participants will upload genuine or manipulated images for analysis
- MediFor analytics and reasoning will authenticate and provide a basis for conclusions
- Participants may respond with a manipulation history for adjudication to document their process if they "Fool MediFor"

Measuring Progress

Program Area	Evaluation	Metric(s)	Goals P1/P2/P3
TA1	**Integrity Analytics**	$P_{Detection}$ and $P_{FalseAlarm}$ for manipulations detected via digital/physical/semantic analytics	Targets Relative to Phase 1 Baseline
TA1, TA2	**Integrity Scores** **Integrity Basis**	Accuracy computed as a function of number, type and severity of manipulations detected Completeness as a percentage basis correct for each manipulation	Targets Relative to Phase 1 Baseline
TA2	**Integrity System**	Throughput of Assets (reports/day) Throughput of analyst reviewed assets (reports/day)	500K generated/instance 100-200 per analyst

MediFor success will ensure novice analysts obtain consistent and complete reports, and experts significantly increase throughput.

Proposal Information

- TA1
  - Proposals may address any or all technologies relevant TA1.1, TA1.2 and TA1.3
  - Only one proposal required per performer
  - Many small awards likely

- TA2
  - Proposers may propose to TA1 as well as TA2***
  - Proposers can team and describe interactions with TA1 performers***
  - Funding for phase 1 is smaller, increasing in phase 2 and 3
  - Proposals will require involvement of cleared personnel
  - One or two awards likely

- TA3
  - Performers may not propose or sub to TA1 or TA2
  - Required data volumes stated in BAA
  - One or two awards likely

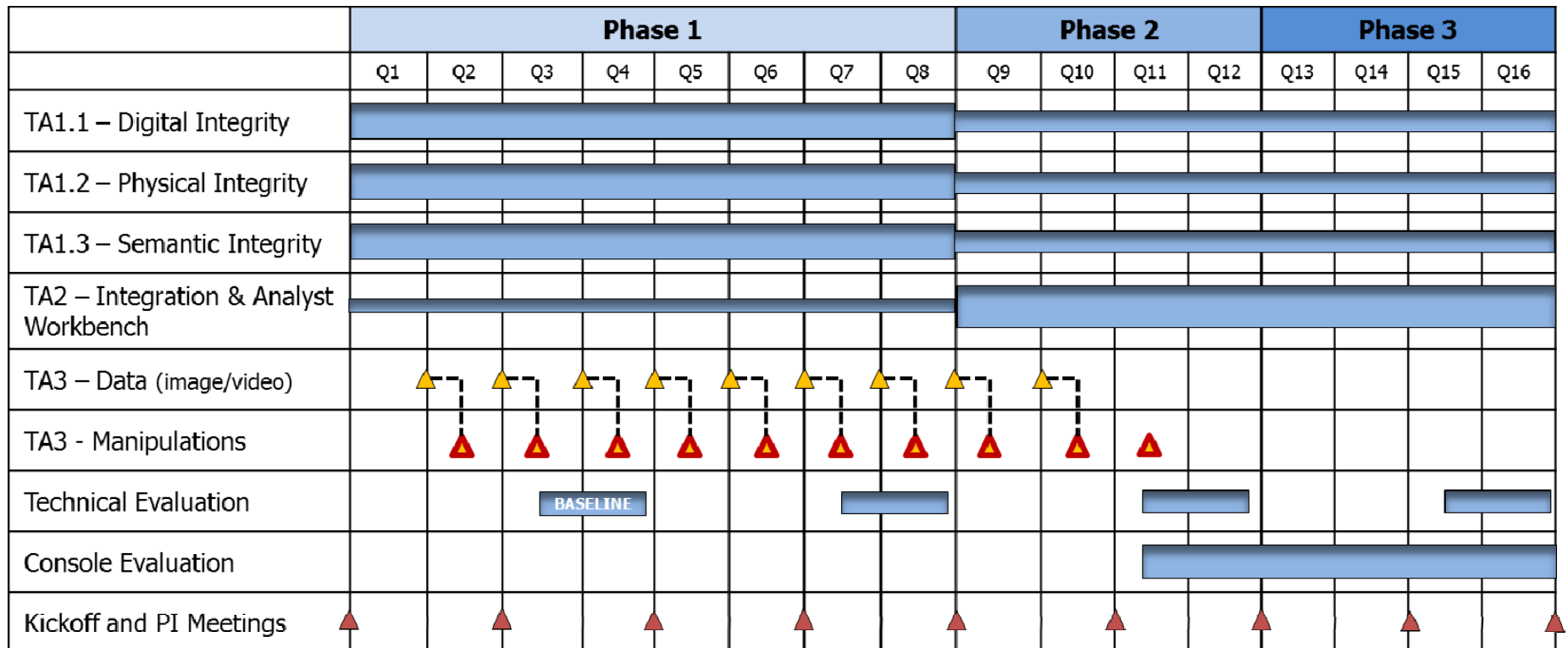***separate proposals are required for TA1 and TA2 tasks

- TA1 – Integrity Analytics Research and Development
  1. Design novel, scalable and automated algorithms robust to expert manipulations
     - TA1.1: Digital Integrity
     - TA1.2: Physical Integrity
     - TA1.3: Sematic Integrity
  2. Report manipulation type, location, parameterization, uncertainty, etc…
  3. Integrate with TA2 teams into the MediFor system

- TA2 - Integrity Reasoning and MediFor Console Development
  1. Integrate TA1 algorithms into a flexible and extensible API
  2. Compile a complete accounting of change and manipulation history
  3. Reason about integrity and produce integrity reports (Score + Basis)
  4. Design a scalable MediFor System (Console and Platform)

- TA3 - Corpora Creation, Manipulation, and Annotations
  1. Collect large dataset of imagery and video with appropriate licenses
  2. Develop adversarial scenarios
  3. Perform manipulations/plant misinformation

	Phase 1								Phase 2				Phase 3			
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16
TA1.1 – Digital Integrity																
TA1.2 – Physical Integrity																
TA1.3 – Semantic Integrity																
TA2 – Integration & Analyst Workbench																
TA3 – Data (image/video)																
TA3 - Manipulations																
Technical Evaluation			BASELINE													
Console Evaluation																
Kickoff and PI Meetings																

Thicker bars indicate a relative increase level of effort

▲ = Meeting Date   △ = 5M Image 50K Videos   △ = 500 Manipulations

Important Points From BAA

- DARPA is **NOT** seeking proposals for the following tasks, technologies, or approaches:
  - Evolutionary or incremental improvements to the existing state of practice
  - Any integrity indicators or reasoning technology primarily reliant on crowdsourcing
  - Program level performance evaluation

Break

- The MediFor Program Q&A session will begin at 1330.

Media Forensics
(MediFor)

Ensuring "seeing is believing"

Dr. David Doermann

I2O

October 2, 2015

Audience Q&A

- **MediFor Program Q&A Session**

Media Forensics
(MediFor)

www.darpa.mil

Ensuring "seeing is believing"